Attorney Docket No. 10559-195001
Serial No. 09/662,679
Amendment dated October 30, 2003
Reply to Office Action dated July 30, 2003

Amendment to the Claims:

This listing of claims replaces all prior versions, and

listings, of claims in the application:

1. (Currently amended)   A method comprising:

receiving audio data having a beat;

forming beat data based on said audio data;

determining a gesture window within which a gesture should

occur, based on a specified time window relative to said beat

data;

playing said audio data and obtaining video data during a

time that said audio data is being played;

segmenting said video data to create a video clip based on

timing data that indicates a specified timing within a gesture

will occur of time including specified timing window; and

automatically determining information related to a gesture

occurring in the video clip only at within the specified timing

window.

2. (Currently amended) The method of claim 1, wherein said

determining includes determining a probability that each of a

plurality of predefined gestures which are performed in the

video clip contains the predefined gesture within the timing

window.

2

3. (Original) The method of claim 2, wherein determining

the probability that the video clip contains each of the

predefined gesture includes evaluations of Hidden Markov Models.

4-6. (Canceled)

7. (Original) The method of claim 1, further comprising

displaying a target gesture to be performed by the subject of

the video data.

8. (Original) The method of claim 1, wherein each video

clip contains video frames.

9. (Currently amended) The method of claim [[1]] 8,

further comprising identifying moving regions in each video

frame in the video clip.

10. (Original) The method of claim 9, further comprising

generating a feature vector for each video frame of the video

clip.

3

11. (Currently Amended) The method of claim 1, further comprising generating a score based on whether the video clip contains ~~the~~ a target gesture.

12. (Original) The method of claim 11, further comprising displaying the score.

13. (Currently amended) The method of claim [[1]] 11, wherein determining if the video clip contains the ~~predefined~~ a target gesture includes generating a gesture probability vector having a plurality of elements, each element being associated with one of a plurality of predefined gestures and representing a probability that the video clip contains each of the associated predefined gestures.

/14. (Currently amended) A system comprising:
    an audio part, receiving audio data having a beat and forming beat data based on said audio data;
    a processor, determining a gesture window within which a gesture should occur, based on a specified time window relative to said beat data;
    a temporal segmentor connected to receive video data during a time that said audio signal is being produced and to create a

4

video clip from the video data ~~based on timing data that~~

~~indicates a specified timing within which a gesture will occur~~

of time including said specified time window; and

a recognition engine, in communication with the temporal

segmentor, to determine if the video clip contains a predefined

gesture, only ~~at~~ within the specified timing window.

15. (Original) The system of claim 14, wherein the

recognition engine includes a plurality of Hidden Markov Models.

16. (Currently amended) The system of claim 14, further

comprising:

~~a timing data source, in communication with the temporal~~

~~segmentor, to provide the timing data to the temporal segmentor;~~

~~and~~

a video source, in communication with the temporal

segmentor, to provide the video data to the temporal segmentor.

17. (Original) The system of claim 14, further comprising a

move subsystem, in communication with the timing data source, to

provide a target gesture to be performed by the subject of the

video data.

5

Attorney Docket No. 10559-195001
Serial No. 09/662,679
Amendment dated October 30, 2003
Reply to Office Action dated July 30, 2003

18. (Original) The system of claim 17, wherein the target gesture is a dance move that is to be performed by the subject of the video data.

19. (Original) The system of claim 17, further comprising a scoring subsystem, in communication with the recognition engine and the move subsystem, to determine if the video clip contains the target gesture.

20. (Original) The system of claim 19, further comprising a display subsystem, in communication with the scoring subsystem, to display a score that is a function of whether the video clip contains the target gesture.

21. (Original) The system of claim 20, wherein the display subsystem is in communication with the move subsystem and is configured to display a gesture request based on the target gesture.

22. (Original) The system of claim 14, wherein the recognition engine is configured to recognize predefined gestures and to produce a gesture probability vector having

6

elements, each element being associated with one of the

predefined gestures and representing the probability that the

video clip contains the associated predefined gesture.

/23-25. (Canceled)

/26. (Currently amended) A computer program product,

tangibly stored on a computer-readable medium, for recognizing

gestures contained in video data, comprising instructions

operable to cause a programmable processor to:

receive audio data having a beat;

form beat data based on said audio data;

determine a gesture window within which a gesture should

occur, based on a specified time window relative to said beat

data;

obtain video data during a time that said audio signal is

being produced;

segment ~~the~~ said video data to create a video clip ~~based on~~

~~timing data that indicates a specified timing within which a~~

~~gesture will occur~~ of the time including said specified timing

window; and

automatically determine if the video clip contains a

predefined gesture within the specified timing window.

7

Attorney Docket No. 10559-195001
Serial No. 09/662,679
Amendment dated October 30, 2003
Reply to Office Action dated July 30, 2003

27. (Canceled)

28. (Currently amended) An audio-visual processing system
including:

a video source to provide video data;

an audio source to provide audio data having a beat;

a speaker to play at least a portion of the audio data; and

a computer program product, tangibly stored on a computer-
readable medium, for recognizing gestures contained in video
data, comprising instructions operable to cause a programmable
processor, in communication with the video source and the audio
source, to:

extract beat data from the audio data;

determine a gesture window within which a gesture
should occur, based on a specified time window relative to said
beat data;

obtain video data during a time that said audio signal is
being produced;

segment the said video data to create a video clip based on
said beat data; and

8

<u>automatically</u> determine if the video clip contains a

predefined gesture within only <u>within</u> a specified timing <u>window</u>

related to said beat data.

29. (Currently amended) The ~~video~~ processing system of

claim 28, wherein the computer program product further includes

instructions operable to cause the programmable processor to:

perform a Hidden Markov Model process to determine if the

video clip contains the predefined gesture.

30. (Currently amended) The ~~video~~ processing system of

claim 28, further comprising a display to display information

based on whether the video clip contains the predefined gesture.